Title

Methylation-Based Deconvolution Unveils Glioblastoma Heterogeneity and Cell-Type Composition Linked to Patient Survival

Aviel Iluz[1,2], Nir Lavi[1,2,3], Hanna Charbit[1,2], Mijal Gutreiman[1,2], Masha Idelson[4], Debora Steiner[4], Etti Ben-Shushan[4], Aviad Zick[5], Amir Eden[6], Anat Mordechai[1,5], Moscovici Samuel[7], Yakov Fellig[8], Alexander Lossos[1,2,5], Joshua Moss[5], Benjamin E. Reubinoff[4], Iris Lavon*[1,2]

[1] Leslie and Michael Gaffin Center for Neuro-Oncology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

[2] Agnes Ginges Center for Human Neurogenetics, Department of Neurology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

[3] Department of Military Medicine and "Tzameret", Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

[4] Hadassah Human Stem Cells Research Center, The Goldyne-Savad Institute of Gene Therapy, and Obstetrics and Gynecology Department, Faculty of Medicine, Hadassah Medical Center, The Hebrew University of Jerusalem, Jerusalem, Israel.

[5] Oncology Department, Sharett institute of oncology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

[6] Department of Genetics, The Hebrew University of Jerusalem, Jerusalem, Israel.

[7] Department of Neurosurgery, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

[8] Department of Pathology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.

* Corresponding author:

Iris Lavon PhD.

Head of Molecular Neuro-Oncology Lab.

Neurology Department and Leslie and Michael Gaffin Center for Neuro-Oncology

Hadassah Hebrew University Medical Center

P.O.Box 12000 Jerusalem 91120 Israel

TeleFax: 972-2-677-7712

Mobile phone: 972-50-8573933

Running title: Glioblastoma Survival Linked to Specific Cell Type Abundance

Abstract

Background:

Glioblastoma is an aggressive and heterogeneous brain tumor, posing significant treatment challenges. DNA methylation profiling enables a detailed characterization of tumor heterogeneity. This study applies methylation-based deconvolution to identify glioblastoma's cellular composition and its association with patient outcomes.

Methods:

Reference atlas with DNA methylation signatures from brain-tumor-relevant cell-types including in vitro-derived human neural progenitors and oligodendrocytes-lineage precursors, was used to deconvolve 263 adult glioblastoma, IDH-Wildtype (IDHwt) samples. Tumor purity was evaluated using RF_Purify, and Kaplan-Meier survival analysis was performed with clinical data from 115 TCGA adult glioblastoma, IDHwt patients.

Results:

Deconvolution revealed distinct glioblastoma cellular compositions consistent with single-cell RNA sequencing studies. Purity analysis distinguished neoplastic (~70%) from non-neoplastic components.

Neoplastic components were primarily oligodendrocyte-lineage cells (69%), including oligodendrocyte-like (42%) and oligodendrocyte precursor-like (27%) populations, along with astrocyte-like (20%) and mesenchymal stem cell-like (11%) signatures. The non-neoplastic fraction comprised macrophages, vascular endothelial cells, and immune cells.

Higher oligodendrocyte-like proportions correlated with poorer survival (MST: 11.9 months; $p<0.011$), while higher astrocyte-like proportions were linked to better survival (MST: 17 months; $p<0.046$). The astrocyte-to-oligodendrocyte ratio emerged as a strong prognostic marker (MST: 16.2 vs. 9 months; $p<0.007$).

Conclusions:

Methylation-based deconvolution effectively characterizes glioblastoma heterogeneity, identifying oligodendrocyte-like predominance and its negative prognostic impact. In contrast, astrocyte-like populations are associated with improved outcomes. The astrocyte-to-oligodendrocyte ratio represents a promising biomarker for patient stratification, potentially guiding personalized treatment strategies.

Keywords

Glioblastoma, methylation-signature, heterogeneity, deconvolution, survival.

Key points

- DNA methylation-based deconvolution reveals glioblastoma cellular heterogeneity.
- Glioblastoma's neoplastic fraction is dominated by oligodendrocyte-lineage cells.
- Higher oligodendrocyte-like levels predict poor survival, while astrocytes improve prognosis.
- Astrocyte-to-oligodendrocyte ratio is a strong prognostic marker for patient survival.

Importance of the study

This study advances neuro-oncology research by using methylation-based deconvolution to uncover the cellular heterogeneity of glioblastoma. By constructing a comprehensive reference atlas and identifying cell-type-specific methylation signatures, it provides new insights into the distinct cellular compositions of glioblastoma subtypes and their prognostic implications. This approach complements single-cell RNA sequencing by minimizing technical challenges of tissue dissociation and offering a detailed view of epigenetic landscapes. Notably, the identification of the astrocyte-to-oligodendrocyte ratio as a strong prognostic marker highlights key cellular determinants of glioblastoma survival. These findings deepen our understanding of glioblastoma biology and offer a practical tool for patient stratification.

Introduction

Glioblastoma, IDH-wildtype (Glioblastoma, IDHwt), is an aggressive and lethal brain cancer, known for its rapid growth and resistance to standard treatments.[1,2] A key feature of glioblastoma is its extensive heterogeneity, which presents challenges for treatment. The tumor contains a complex mix of neoplastic and non-neoplastic cells, including glioma initiating cells, immune cells, endothelial cells, and glial cells.[2-5]

Genome-wide DNA methylation profiling has become a transformative tool for classifying CNS tumors, enabling patient stratification beyond traditional diagnostics.[6,7] In glioblastoma, while DNA methylation profiling is not required for diagnosis, it helps identify distinct subclasses, including receptor tyrosine kinase I (RTK I), receptor tyrosine kinase II (RTK II), and mesenchymal (MES).[6,7] However, a key challenge in glioblastoma classification is its inherent heterogeneity, which complicates subgroup assignment and is linked to variations in patient prognosis.[8,9] Studies of newly diagnosed glioblastomas emphasize the spatial diversity within these methylation subclasses, highlighting the complexity of classification.[2,10] This intratumoral heterogeneity is increasingly recognized as a major factor in treatment resistance and disease recurrence.[11]

Understanding the cellular drivers of glioblastoma heterogeneity is essential for advancing research and developing effective therapies. Various methods, including proteomic and transcriptomic analyses, have been used to explore this complexity.[2-4,12] Recent breakthroughs in single-cell RNA sequencing (scRNA-seq) have provided new insights into the cellular diversity of glioblastoma, revealing distinct subpopulations of neoplastic cells and their interactions with the tumor microenvironment [13].[2-5] A landmark study[3] identified four distinct cellular states within the neoplastic fraction of glioblastoma: astrocyte-like (AC-like), oligodendrocyte progenitor cell-like (OPC-like), mesenchymal-like (MES-like), and neural progenitor cell-like (NPC-like). These findings underscore the tumor's complex differentiation and plasticity.

An alternative approach to investigating glioblastoma cellular composition, in addition to scRNA-seq, involves analyzing DNA methylation signatures.[14] Each cell type has a unique methylation signature that reflects its tissue identity.[15,16] Tumors also exhibit stable methylation signatures, even after oncogenic events, enabling precise sub-classification of brain tumors.[17,18] These signatures have been validated in thousands of CNS tumors, facilitating accurate

diagnoses and better treatment guidance.[6,19-21] This method was incorporated into the 2021 WHO Classification of CNS Tumors.[21]

The DNA methylation signature of a tumor reflects the cumulative methylation patterns of all its constituent cell types, including both normal and neoplastic cells. This composite profile captures the cellular admixture within the tumor, providing insights into its complex cellular makeup.[14] Methylation-based deconvolution utilizes a reference atlas of normal cell type methylation signatures to estimate tissue composition and has been used to analyze tissue contributions in circulating cell-free DNA (cfDNA).[22,23] A similar deconvolution methodology has been applied to investigate the composition of non-neoplastic immune cells within the glioblastoma TME.[14]

To the best of our knowledge, methylation signatures have not been specifically used to investigate the neoplastic component of brain tumors. In our study we take a comprehensive approach to analyze both neoplastic and non-neoplastic components of glioblastoma tumors using methylation-based deconvolution. By examining a broad set of methylation signatures from various brain tumors,[6] with a particular focus on glioblastoma, we aim to elucidate glioblastoma heterogeneity across all cell types, with a specific emphasis on the neoplastic fraction. This analysis may also shed light on the potential cell of origin of glioblastoma, a critical aspect for understanding the initial events in tumorigenesis and identifying early intervention targets.[24]

To validate and contextualize our findings, we will compare our results with data obtained through complementary methodologies, including scRNA-seq and the deconvolution of immune cell profiles from bulk glioma DNA methylation data.[3,14,25-27] This integrative approach is expected to enhance our understanding of glioblastoma's cellular complexity and inform future prognostic and therapeutic strategies.

Materials and Methods

Generation of human NP and oligodendrocyte lineage precursor cells

Highly enriched cultures of neural progenitors [28] were derived from human embryonic stem cells (hESCs) using an established protocol (designated in this study as "NP").[29] Briefly, hESCs (HAD-C100[30]) at passages 22 to 28 were cultured on human recombinant Laminin 521 (Biolamina, Sundbyberg, Sweden) in NutriStem medium supplemented with growth factors (Biological Industries, Beit Haemek, Israel) and 50 units/mL penicillin and 50 µg/mL streptomycin (Pen-Strep, Thermo Fisher Scientific, Waltham, MA, USA). Oligodendroglial cells at various developmental stages were derived from these enriched NP cultures as previously described.[29] Briefly, NP were first guided to differentiate into OLIG2-positive progenitor cells (designated as "OLIG2") and subsequently directed toward the oligodendroglial lineage. This process resulted in an enriched population of Olig2+/Nkx2.2+ oligodendrocyte progenitor cells (designated as "OPC"), which were further differentiated into O4+ pre-oligodendrocytes (designated as "pre-oligodendrocytes").

DNA methylation profiling

Genome-wide DNA methylation profiling of NP, OLIG2, OPC, and pre-oligodendrocytes was performed using the Illumina EPIC array, with data generated at the Genomics and Proteomics Core Facility at the German Cancer Research Center (DKFZ) in Heidelberg, Germany. Data processing, filtering, and normalization were carried out using the *minfi* R package (v1.34.0).

Data collection

We collected 450K or EPIC Illumina methylation signature data to assemble the various sample sets analyzed in this study. All data were processed as beta-value CSV files. To ensure consistency, EPIC beta-value CSV files were intersected with the 450K methylated cytosine-guanine dinucleotides (CpG) probe set, retaining only common probes for downstream analysis.

Tumor sample sets for deconvolution and survival analysis

The adult IDH-wildtype glioblastoma dataset was sourced from the Heidelberg reference set[6] and included 263 cases classified according to the 2021 WHO Classification of CNS Tumors. These cases were further categorized into GBM_RTK_I (n=64), GBM_RTK_II (n=143), and GBM_MES (n=56). Beta-values of CpG sites were extracted from the nanoDx pipeline dataset (capper_et_al.h5)[31,32] as CSV files using the *rhdf5* R package (v2.32.4).[33]

Glioma-initiating cell (GIC) sample set (n=20) was obtained as beta-values CSV files from a publicly available dataset referenced in Vinel et al.[34]

Non-glial tumors set obtained from TCGA database as beta-values CSV files and included data of bladder urothelial carcinoma (BLCA) (n=292), breast carcinoma (n=721), kidney renal papillary cell carcinoma (KIRP) (n=210) and prostate adenocarcinoma (PRAD) (n=127).

For survival analysis, we used adult IDH-wildtype glioblastoma cases, with available clinical information from the TCGA cohort, classified according to the 2021 WHO classification of CNS Tumors. The cohort of the study included 115 samples classified into GBM_RTK_I (n=24), GBM_RTK_II (n=56), and GBM_MES (n=35) subclasses, using the DKFZ classifier[6] (v11b4) on the Molecular Neuropathology platform (www.molecularneuropathology.org).

The majority of patients in this cohort were treated according to the Stupp protocol, which includes radiation therapy with concurrent and adjuvant temozolomide.

Bioinformatics and statistical analysis

Constructing reference atlases

Reference atlases were constructed using the CpG feature selection method from Moss et al.[23] The top 100 uniquely hypermethylated and 100 uniquely hypomethylated CpGs were selected for each cell type, including neighboring CpGs within 50 base pairs, to generate tissue-specific CpG sites for each atlas.

Glioblastoma tumors reference atlases: We used two reference atlases for glioblastoma deconvolution analysis. The initial reference atlas included 16 cell type signatures: B-cells, CD4 T-cells, CD8 T-cells, NK cells, neutrophils, vascular endothelial cells, monocytes, microglia, MSC, cortical neurons, astrocytes, oligodendrocytes, NP, OLIG2, OPC, and pre-oligodendrocytes, covering a total of 4,712 CpG sites. Based on the results from deconvolving Heidelberg glioblastoma samples, we refined this atlas to include 14 cell type signatures by merging OLIG2, OPC, and pre-oligodendrocytes into a single oligodendrocyte-lineage precursor component (designated as "OP"), reducing the total to 4,111 CpG sites. This refined atlas was subsequently used to deconvolve Heidelberg glioblastoma, TCGA and GEO glioblastoma, and GIC samples.

Non-glial tumors reference atlas: Used to deconvolve the methylation profiles of non-glial tumors from the TCGA database. It includes 25 cell type signatures sourced directly from publicly available data resources by Moss et al.[23], comprising a total of 7,390 CpG sites

Validation reference atlas for normal cell types: Used to deconvolve the methylation profiles of normal cell type signatures. It includes 11 cell type signatures sourced directly from publicly available data by Moss et al.[23], with N-1 samples used to construct the atlas and one sample excluded for validation. It comprises a total of 3,274 CpG sites.

Deconvolution of Methylation Signatures Using NNLS Linear Regression

Non-negative least squares (NNLS) linear regression was used to deconvolve the beta-values of methylation signatures into cell-type components, utilizing the relevant reference atlas, as outlined by Moss et al.[23]

DNA tumor purity

Tumor purity was calculated using the *RF_purify* R package (v0.1.2) with the ABSOLUTE method. This method analyzes Illumina 450K methylation data to estimate purity by applying a Random Forest machine learning algorithm, trained on copy number variations (CNVs) inferred from methylation patterns.[35]

Survival analysis

Kaplan–Meier analysis and log-rank tests were conducted on 115 TCGA glioblastoma samples using the *survminer* (v0.5.0) and *survival* (v3.7-0) R packages ().[36] An optimal cutoff was determined to classify samples into high and low cell type proportion groups, each exhibiting distinct overall survival probabilities [17], as previously described.[37] P-values were calculated, with a significance threshold set at $p < 0.05$.

Statistical Analysis

To assess statistically significant differences in cell type proportions among glioblastoma subclasses, we conducted pairwise Student's t-tests for each subclass pair.

Results

Identification of Cell Type Composition and Proportions in IDH-Wildtype Glioblastoma and Its Subtypes: GBM_RTK_I, GBM_RTK_II, and GBM_MES

To investigate cell type composition in glioblastoma, we constructed a reference atlas of methylation signatures using data from 16 distinct cell types. These included neurons, astrocytes, oligodendrocytes, immune cells (B cells, CD4+ T cells, NK cells, CD8+ T cells, and neutrophils), monocytes, microglia, and neural progenitors[28], which are multipotent cells capable of differentiating into various neural cell types. Additionally, the atlas captured three stages within the oligodendrocyte lineage, representing a continuum of differentiation: OLIG2-positive progenitors (expressing the transcription factor OLIG2 essential for oligodendrocyte development), oligodendrocyte progenitor cells (OPCs), and pre-oligodendrocytes marked by O4 expression. This atlas was generated by selecting distinct differentially methylated cytosine-guanine dinucleotides (CpGs) for each cell type, following previously described and validated methodologies.[23] The cell type-specific methylation signatures are crucial for deconvoluting the cellular composition of brain tumor samples, enabling the quantification of cell type proportions. Our reference atlas comprises approximately 4,700 CpGs across these cell types (**Figure S1**).

To estimate the cellular composition of bulk glioblastoma tumor samples, we applied a deconvolution method adapted from cfDNA tissue-origin studies.[23] Using 450K methylation data, we analyzed 263 adult IDH-wildtype glioblastoma samples, classified according to their methylation profiles as previously described.[6]

In glioblastoma samples (**Figure S2**) glial cells constituted the majority (62%) of the cellular composition, including contributions from oligodendrocytes-like signature (26.3%), astrocytes-like (13.4%), and pre-oligodendrocytes-like (O4+, 22%). Cortical neurons signature was minimally represented (2.6%). Additionally, glioblastoma displayed 7.3% mesenchymal stem cells-like (MSC), 7.8% tumor associated macrophages [3] comprising microglia (3.6%) and monocytes (4.2%), 13% vascular endothelial cells and 7.2% immune cells (B-cells, CD4 T-cells, NK cells, CD8 T-cells, and neutrophils).

To further refine the analysis of glioblastoma cell type composition, OLIG2, OPC, and pre-oligodendrocytes were merged into a single component termed oligodendrocyte-lineage precursor (OP). This adjustment was necessary because OLIG2 and OPC contributed

minimally to glioblastoma proportions (average 0.3 ± 1.2%). However, their integration ensured representation while preserving the accuracy of cell type classification. As a result, an alternative reference atlas was established, encompassing 14 distinct cell types (**Figure 1**).

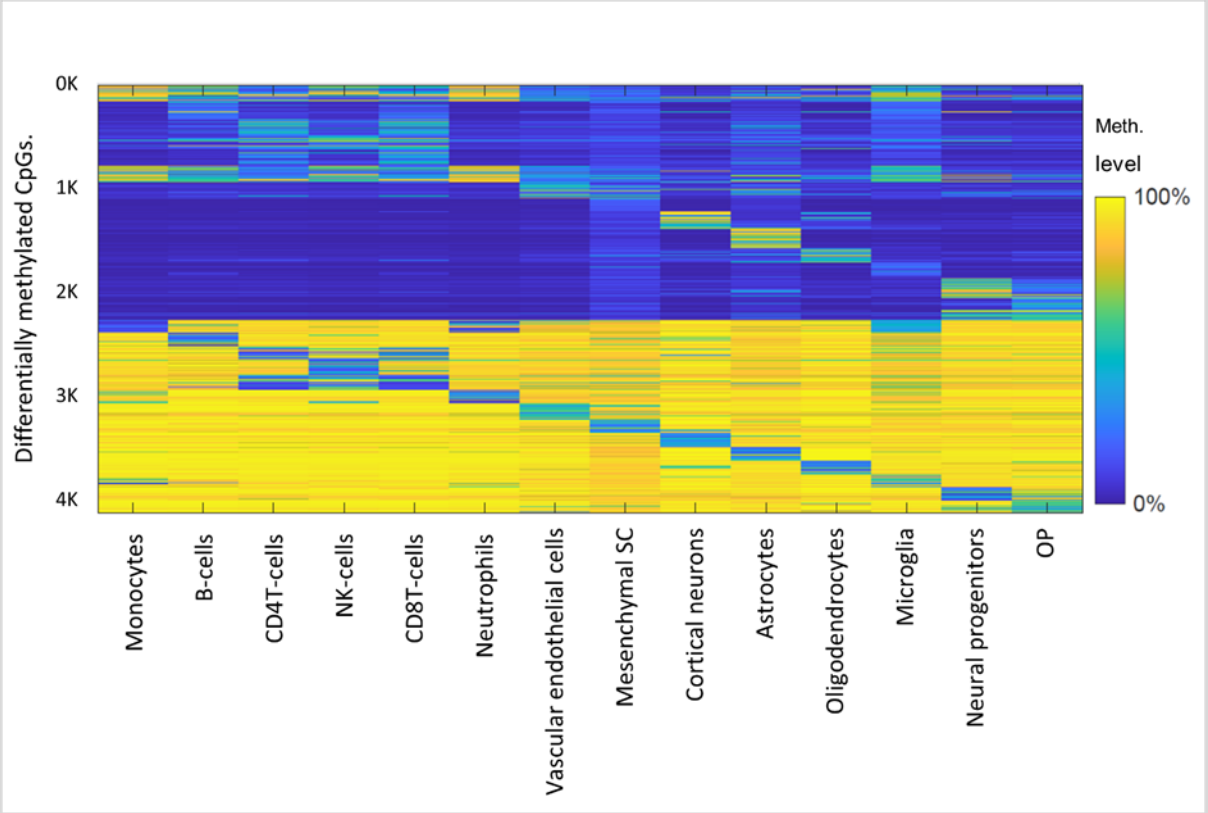

Figure 1. Reference atlas for glioma deconvolution and cell-type identification using methylation signatures.

Methylation atlas comprises 14 tissues/cell types (columns) across 4712 CpGs sites (rows) that are located in 2817 genomic blocks of 500 bp each. Feature selection was carried out as described in a previous study.[23] Specifically, the top 100 uniquely hypermethylated CpGs and hypomethylated CpGs for each cell type, yielding a total of 3200 tissue-specific CpGs, with neighboring CpGs within 50 bp were also included.

This adjustment had a minimal impact on the estimated proportions of astrocytes, cortical neurons, MSCs, TAMs, vascular endothelial cells, and immune cells. However, a slight difference was observed when comparing the separated oligodendrocyte lineage precursors to the combined OP component, the proportion shifted from 26.3% to 29.5%. Specifically, pre-oligodendrocytes accounted for 22% in the separated samples, whereas the combined OP component accounted for 18.9% (**Figures 2, S2**).

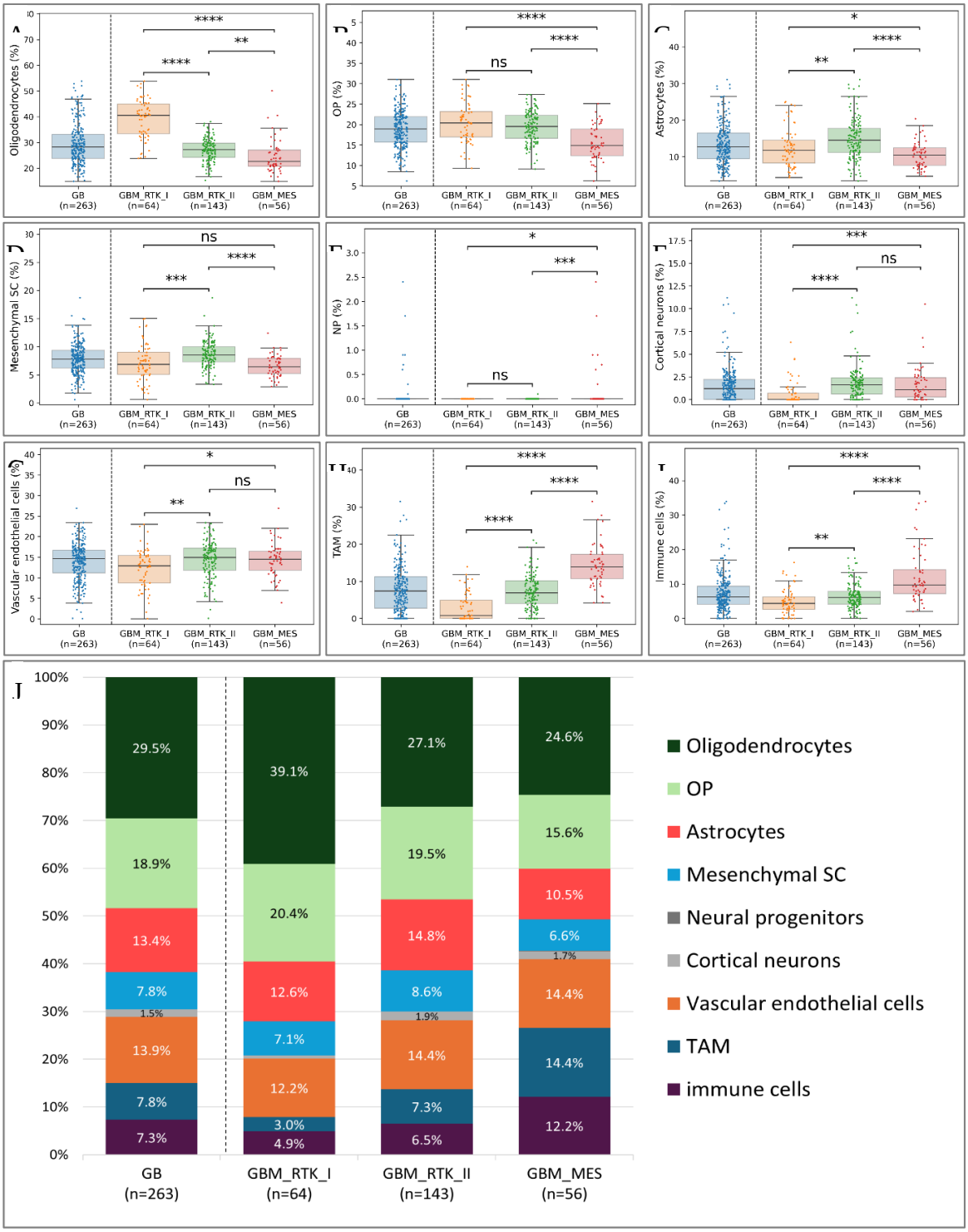

Figure 2. Cell-type proportions derived from deconvolution analysis in glioblastoma IDHwt and their differences across glioblastoma subtypes.

(A-I) Boxplots showing the proportions of cell types derived from deconvolution analysis of methylation data across glioblastoma GB, (n=263, blue); GBM RTK_I (n=64, orange), GBM RTK_II (n=143, green), and GBM_MES (n=56, pink), The central line represents the median, the edges of the boxes indicate the interquartile

range (IQR), and the whiskers extend to 1.5 times the IQR. Plots were generated using Python. (J) Bar plot summarizing the proportions of normal cell types derived from the deconvolution of methylation profiles of 263 glioblastoma IDHwt samples (leftmost bar) classified as GBM RTK_I (n=64), GBM RTK_II (n=143), and GBM_MES (n=56) based on the atlas of 14 cell types. The average percentage for each cell type is displayed within its respective bar. Immune cell types (B cells, CD4 T cells, NK cells, CD8 T cells, and neutrophils) are grouped as "immune cells", and microglia and monocytes as grouped as "TAM".

The boxplots display the cell type proportions for (A) oligodendrocytes (B) OP (C) astrocytes (D) Mesenchymal SC [36] NP [24] Cortical neurons [36] Vascular endothelial cells [24] TAM (I) Immune cells. Black lines indicate significant differences between groups, as determined by pairwise Student t-test (*p<0.05; **p<0.01; ****p<0.0001; ns: not significant). Plots generated using Python. (OP, oligodendrocyte-lineage precursors; NP, Neural progenitors)

Our findings revealed significant variation in cell type proportions among glioblastoma subclasses. Specifically, astrocyte proportions differed significantly between GBM_RTK_I and GBM_RTK_II (12.6% vs. 14.8%, p<0.01) and between GBM_RTK_II and GBM_MES (14.8% vs. 10.5%, p<1.60E-07), but not between GBM_RTK_I and GBM_MES. Oligodendrocyte proportions also showed significant pairwise differences in all comparisons: GBM_RTK_I vs. GBM_RTK_II (39% vs. 27%, p<1.012E-32), GBM_RTK_I vs. GBM_MES (39% vs. 24.6%, p<2.39E-20), and GBM_RTK_II vs. GBM_MES (27% vs. 24.6%, p<1.85E-03). Additionally, significant differences were observed in the proportions of other cell types, including OP, MSC, immune, and TAM cell populations. All p-values were calculated using a two-tailed t-test. **(Figure 2)**.

Validation of Deconvolution Method Through Analysis of Non-Glial Tumors and Normal Cell Types

To validate the methylation-based deconvolution method, we applied it to methylation data from non-glial tumors, including bladder urothelial carcinoma (BLCA), breast carcinoma, kidney renal papillary cell carcinoma (KIRP), and prostate adenocarcinoma (PRAD). The reference methylation atlas, comprising 25 cell types, was sourced from a previous study.[23] Analysis showed that the highest proportion in each tumor type matched the normal cell type of its tissue origin: bladder cells (74%) in BLCA, breast cells (62%) in breast cancer, kidney cells (79%) in KIRP, and prostate cells (90%) in PRAD (**Figure S3**).

To further validate our deconvolution method, we applied it to 11 samples from various normal cell types, using a reference atlas derived from methylation data of the same cell types. The

analysis showed that over 90% of each sample matched the expected normal cell type, e.g., 100% accuracy in the cortical neuron sample (**Figure S4**).

Assessing the Neoplastic and Non-Neoplastic Components of Glioblastoma Using RF_Purify Method

Using cell type proportions from the deconvolution analysis, we aimed to distinguish between neoplastic and non-neoplastic components in glioblastoma tumors. This was done using the RF_Purify method, which applies the ABSOLUTE score to infer tumor purity.[38] We correlated purity scores with the proportions of seven cell types identified in the deconvolution analysis, excluding cortical neurons (2%) and NP (0%) due to their minimal contribution. The analyzed components included immune cells, TAM, vascular endothelial cells, OP, MSC, astrocytes, and oligodendrocytes (**Figure 3**).

Significant negative correlations were observed between tumor purity and the proportions of immune cells (R=-0.60, p<1.E-26) and TAM (R=-0.76, p<3.E-51), with vascular endothelial cells showing a weaker negative correlation (R=-0.20, p<1.E-03) (Figures 3B, 3D). In contrast, tumor purity correlated positively with glial lineage cells, including OP (R=0.68, p<1.E-37), astrocytes (R=0.33, p<6.E-08), oligodendrocytes (R=0.41, p<6.E-12), and MSC (R=0.22, p<4.E-04) (Figures 3A, 3D). When the proportions of all positively correlated components (glial cell types and MSC) were combined, the correlation with tumor purity was even stronger (R=0.79, p<2.E-56), suggesting these components contribute more substantially to the neoplastic portion of the tumor than individually. All p-values were calculated using a two-tailed correlation t-test. (**Figures 3C, 3D**).

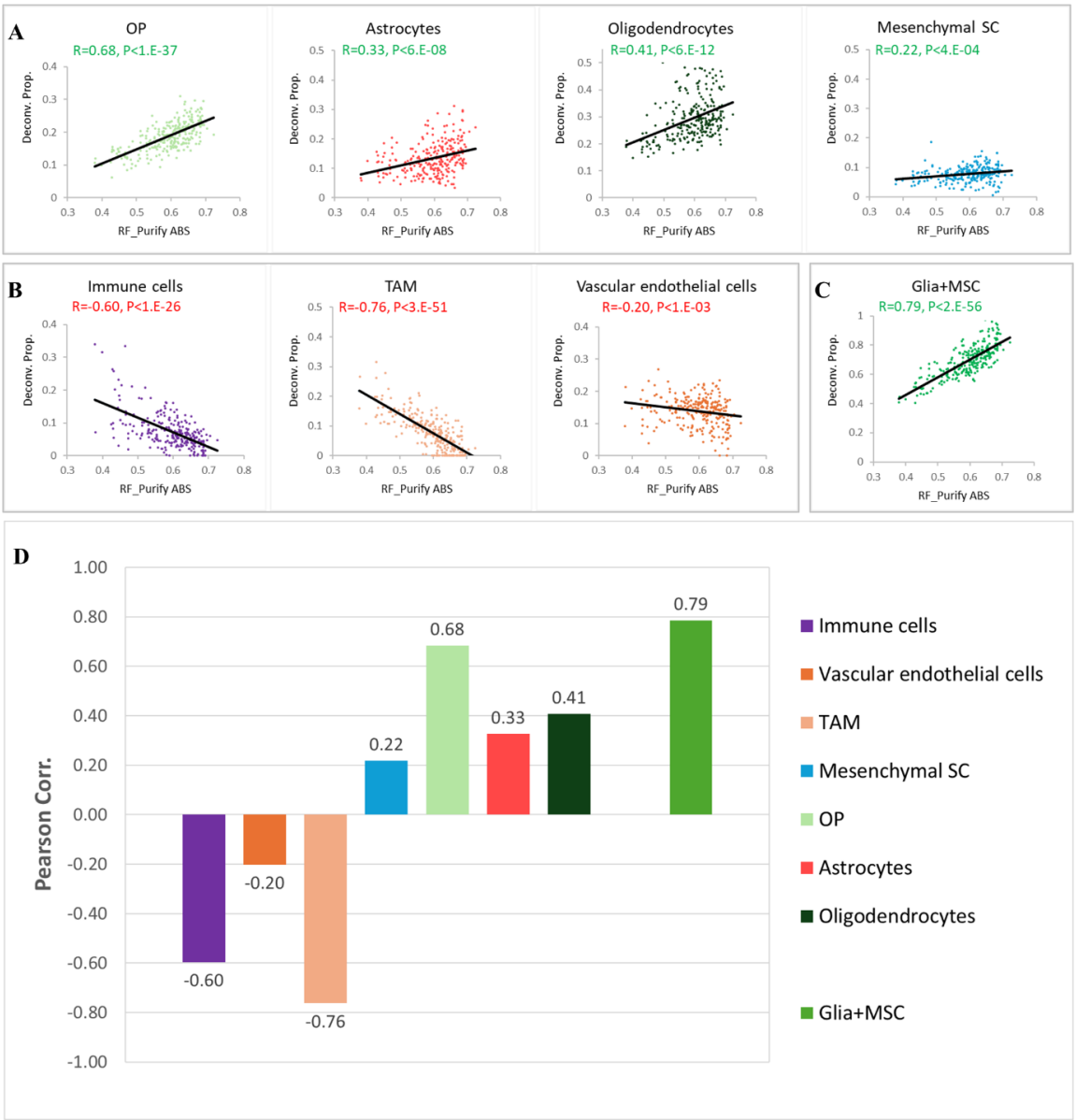

Figure 3. Correlation between cell-type proportions and purity estimates calculated by RF_Purify_ABSOLUTE.

Correlation between cell type proportions derived from the deconvolution of all 263 glioblastoma samples and sample purity estimated by the RF_Purify algorithm using the ABSOLUTE method. The analyzed cell types include (A) immune cell totals (purple), vascular endothelial cells (dark orange), TAM (light orange), (B) mesenchymal stem cells (MSC; blue), three glial cell types – oligodendrocyte-lineage precursors (OP; light green), astrocytes (red), and oligodendrocytes (dark green); and (C) the combined total of the four positively correlated types (glia+MSC, green). (D) An overview histogram illustrates the correlation values between proportions for eight cellular components (listed on the right) and sample purity.

Comparison of Methylation-Based Deconvolution Results with scRNA-Seq Studies

Our study demonstrates that the neoplastic fraction, derived from purity analysis of 263 samples, has an average of 70% (±11%). This result aligns closely with estimates from five

previous studies, four based on scRNA-Seq[3,12,25-27] and one using methylation-based deconvolution,[14] which reported an average neoplastic fraction using CNV-based methods at 72% (ranging from 52% to 87%) (**Table 1, Figure 4**). The average of the total proportion of immune cells in our study was 7.3% (±5.4%), higher than the 3.3% (ranging from 0.5% to 7.2%) reported across four other studies. Specifically, we observed: B-cells, 3.4% (±1.5%) compared to 1.1% reported only in one of the five studies; CD4 T-cells, 0.3% (±0.8%) compared to 0.2% from one study; NK cells, 0.4% (±0.7%) versus 0.6% (range 0.0% to 1.4%) across three studies; CD8 T-cells, 1.1% (±1.4%) compared to 1.5% (range 0.5% to 3%) from four studies; Neutrophils, 2.2% (±4.6%) compared to 1.3% (range 0.1% to 2.8%) from three studies. TAM, accounted for an average of 7.8% (±6%), compared to 5.4% (range 0.2% to 7.7%) across five studies. Cortical neurons showed an average of 1.5% (±1.8%) compared to 0.5% (range 0% to 1.4%) in three studies (**Table 1, Figure 4**). A notable difference emerged for vascular endothelial cells, which had an average proportion of 13.9% (±4.5%) in our study compared to the 1.7% (range 0.3% to 2.8%) in four other studies (data not shown).[12,14,25-27]

Table 1: Comparison of Methylation-Based Deconvolution Results with other studies.

TME fractions	cell types	Neftel	Yu	Darmanis	Yuan	Singh	other studies (avg.)	our study (avg.)
Non-Neoplastic fraction	B cells	-	-	-	-	1.1%	1.1%	3.4%
	CD4T cells	-	-	-	-	0.2%	0.2%	0.3%
	NK-cells	0.0%	1.4%	-	-	0.4%	0.6%	0.4%
	CD8T cells	1.2%	3.0%	-	0.5%	1.3%	1.5%	1.1%
	Neutrophils	0.1%	2.8%	-	-	1.1%	1.3%	2.2%
	TAM	5.4%	6.4%	7.7%	0.2%	7.5%	5.4%	7.8%
	Cortical neurons	-	-	1.4%	0.0%	0.2%	0.5%	1.5%
	Normal Oligodendrocytes	2.8%	9.1%	6.9%	1.6%	6.2%	5.3%	-
	Normal Astrocytes	-	0.6%	1.4%	-	-	1.0%	-
Neoplastic fraction		87%	52%	59%	83%	78%	72%	70%

Summary of the proportions of various cell types and neoplastic fraction identified in our study compared to those reported in other studies and their average. In the right section, four columns represent the scRNA-Seq-based studies: Neftel et al.[3], Yu et al.,[27] Darmanis et al.[25] and Yuan et al.[26] The fifth column on the right displays findings from methylation-based analysis (Singh et al.[14]). (TME, tumor-microenvironment)

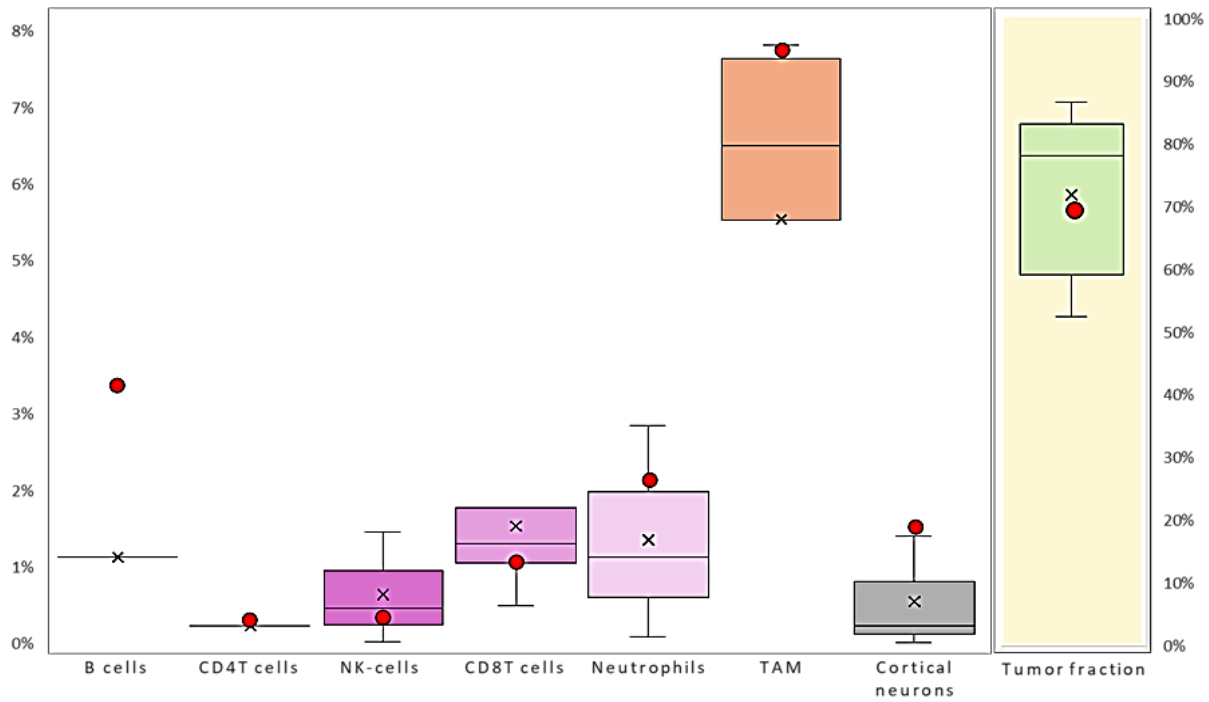

Figure 4: Comparison of proportions of different cell types between our study and other studies.

Boxplots illustrating the distribution of cell type proportions and the neoplastic fraction in glioblastoma samples from other studies, with each boxplot representing a specific cell type. The boxplot on the far right depicts the neoplastic fraction on a 100% scale. Our findings are indicated by a red circle for comparison. The central line in each boxplot shows the median, the X marks the mean, the box edges indicate the interquartile range (IQR), and the whiskers extend to 1.5 times the IQR.

Validation of Purity Analysis: Deconvolution Results in Bulk Glioblastoma Tumors Versus Glioma-Initiating Cells

To validate our purity analysis that links the three glial cell types and the MSCs to the neoplastic portion of glioblastoma, we applied the analysis to methylation signatures from CD133 (Prominin-1)-enriched cells isolated from glioblastoma bulk tumors.[34] CD133, a well-established marker for cancer stem cells, selects for glioma-initiating cells (GICs), which are known to be enriched for neoplastic cells. These GICs were derived from glioblastoma bulk tumors representing the GBM_RTK_I, GBM_RTK_II, and GBM_MES subclasses, providing a relevant model for assessing whether our approach effectively captures the neoplastic cell population within the tumor.

As expected, deconvolution of GICs revealed a significantly higher tumor fraction composed of three glial lineage cell types and MSCs compared to glioblastoma bulk tumors, with 88% in

GICs versus 70% in glioblastoma ($p<2.2E-20$, two-tailed t-test, Figure S5). Specifically, GICs exhibited significantly increased proportions of astrocytes (19% vs. 13%, $p<8.7E-04$), OP (26% vs. 19%, $p<3.1E-05$), and MSCs (16% vs. 8%, $p<5.6E-11$), while oligodendrocyte proportions remained similar (28% vs. 30%, $p<0.21$). In contrast, vascular endothelial cells, immune cells, and TAMs were significantly reduced in GICs compared to glioblastoma bulk tumors (vascular endothelial cells: 7% vs. 14%, $p<2.4E-09$; immune cells: 3% vs. 7%, $p<1.1E-07$; TAMs: 0% vs. 8%, $p<6.5E-54$) (Figure S5). All p-values were calculated using a two-tailed t-test. These findings demonstrate that cell types positively correlated with tumor purity in glioblastoma (Figure 3A) are more abundant in GICs, whereas those negatively correlated with tumor purity (Figure 3B) are present at lower proportions.

Correlation Between Cell Type Proportions and Patient Survival: Prognostic Significance of the Astrocyte-to-Oligodendrocyte Ratio and Its Association with Glioblastoma Subclasses

We explored the relationship between patient survival and cell type proportions using Kaplan-Meier analysis on 115 TCGA glioblastoma samples with clinical data (**Figure 5**).

Our analysis revealed significant associations between survival and the proportions of oligodendrocytes and astrocytes, while OP and microglia showed trends. Higher oligodendrocyte proportions were linked to reduced OS, with a median survival time (MST) of 11.9 months compared to 17.4 months, resulting in a median survival time difference (MSTD) of 5.5 months ($p<0.011$, log-rank test, **Figure 5A**). OP showed a similar but nonsignificant trend (MSTD=1.4 months, $p<0.157$, log-rank test, **Figure 5B**).

Conversely, higher astrocyte proportions correlated with improved OS, with an MST of 17 months compared to 11.4 months (MSTD=5.6 months, $p<0.046$, log-rank test**, Figure 5C**). Microglia also trended toward higher OS but did not reach significance (MSTD=5.3 months, $p<0.07$, log-rank test, **Figure 5D**). No significant correlations were found for other cell types (data not shown).

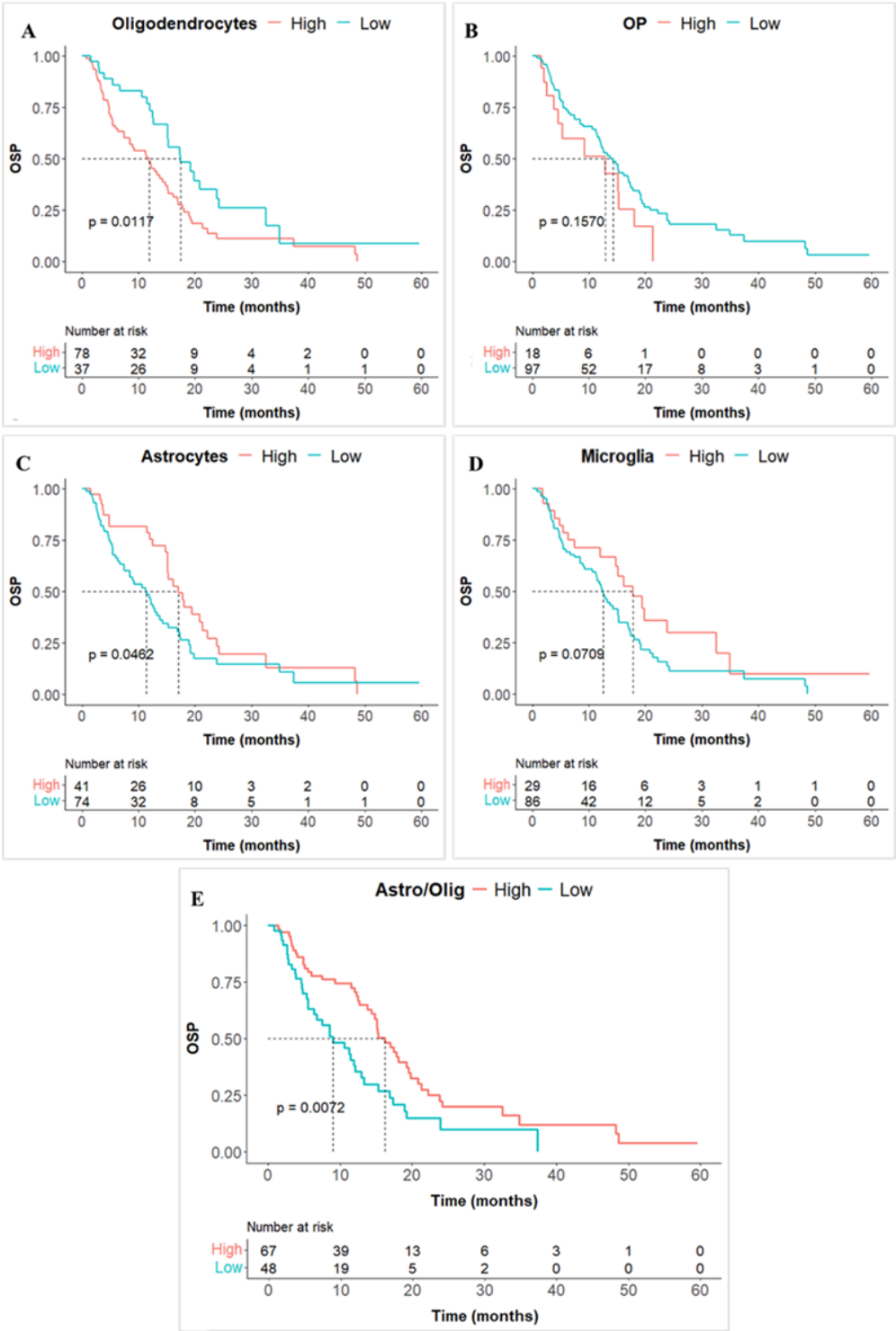

Figure 5: Kaplan-Meier plots based on cell type proportions.

Kaplan-Meier survival analysis shows OSP differences based on high and low proportions of specific cell types. Each plot displays the range of proportions and the cutoff value, determined as the point that best separates OSP outcomes for the two groups: (A) Oligodendrocytes (cutoff=25.6%), (B) Oligodendrocyte-lineage precursors (OP;

cutoff=23.6%), (C) Astrocytes (cutoff=9.8%), and (D) Microglia (MCG; cutoff=3.7%), (E) Astro/Olig proportion ratio (cutoff ratio=0.238).

Building on the contrasting effects of astrocyte and oligodendrocyte abundance on OSP, we examined the astrocyte-to-oligodendrocyte (Astro/Olig) ratio's impact on survival. A high Astro/Olig ratio (>0.238) was significantly associated with higher OSP (MST=16.2 months), while a low ratio (<0.238) corresponded to reduced OSP (MST=9 months), with an MSTD of 7.2 months (p<0.007, log-rank test, **Figure 5E**).

Given the Astro/Olig ratio's impact on OSP and the variations in astrocyte and oligodendrocyte proportions among glioblastoma subclasses (Figure 2), we explored their association with survival outcomes. GBM_RTK_II samples were predominantly linked to long-term survival (LTS, 84%), GBM_RTK_I to short-term survival (STS, 75%), and GBM_MES showed a balanced distribution (LTS 40%, STS 60%). This subclass distribution between LTS and STS groups was statistically significant (p<1.96E-07, chi-square test). (**Figure S6**).

However, when survival analysis was based solely on glioblastoma subclass without considering cell type proportions, no significant differences in OSP were observed (**Figure S7**).

Discussion

Our study establishes methylation-based deconvolution as a powerful tool for unraveling glioblastoma's cellular composition, shedding light on tumor heterogeneity and linking cell type proportions to patient survival.

By leveraging a reference atlas of methylation signatures from glioblastoma microenvironment cell types, including in vitro-differentiated human oligodendrocyte-lineage precursor (OP) and Neural progenitor,[28] we identified key cellular components within the tumor. Notably, glial cells dominated the glioblastoma composition (62%), primarily oligodendrocytes-like, oligodendrocyte precursor-like and astrocytes-like signatures, while cortical neurons contributed minimally. Additionally, we identified presence of mesenchymal stem cells (MSC)-like, vascular endothelial cells, tumor-associated macrophages, and immune cells, further underscoring the complexity of the glioblastoma microenvironment.

Our results were consistent with prior deconvolution analyses,[3,12,14,25-27] with the notable exception of a higher vascular endothelial cell proportion (13%), consistent with a recent report

indicating 9.9% vascularization in glioblastoma.[39] This finding underscores the well-documented angiogenic nature of glioblastoma and reinforces the reliability of our approach in capturing glioblastoma's vascular complexity.

To distinguish neoplastic from non-neoplastic components, we integrated RF_Purify and ABSOLUTE score analyses**,** confirming that oligodendrocyte-lineage precursors, astrocytes, oligodendrocytes, and MSCs, constitute the neoplastic fraction**,** while immune cells, TAM, and vascular endothelial cells predominantly belong to the tumor microenvironment. The combined neoplastic fraction averaged 70%, consistent with prior estimates (~72%)[3,12,14,25-27].

Aligned with prior research,[3,39] our findings indicate that astrocytes-like and oligodendrocytes-like signatures constitute a significant part of the glioblastoma neoplastic fraction. The consistently reported low abundance of normal astrocytes and oligodendrocytes in glioblastoma samples[3,12,14,25-27] supports the conclusion that elevated proportions we detected primarily reflect the neoplastic component.

We propose that our astrocyte and MSC methylation signatures correspond to the AC-like and MES-like states identified by scRNA-seq[3-5] as part of the neoplastic fraction of glioblastoma. However, assigning a methylation signature of oligodendrocyte-lineage precursor (OP) and mature oligodendrocytes to the less differentiated NPC-like and OPC-like states, is more challenging. Our analysis shows that early oligodendrocyte lineage stages (OLIG2^{+} or OPC-like cells) contribute minimally to the OP component. Instead, most oligodendrocyte-lineage cells exhibit advanced differentiation, as indicated by pre-oligodendrocyte (O4-positive) and mature oligodendrocyte methylation signatures. The discrepancy between our findings and previous studies[3-5] likely arises from methodological differences between single-cell transcriptional profiling and bulk methylation analysis. Further studies, particularly single-cell methylation analysis, are needed to validate this interpretation. Despite these methodological differences, both approaches suggest that these populations represent stages within the oligodendrocyte differentiation lineage.

We show that the oligodendrocyte differentiation lineage constitutes approximately 50% of the bulk tumor and 71% of the neoplastic fraction. Our findings build upon prior reports by Liu et al.[40], Persson et al.[41], and Ligon et al.[42], which demonstrated that oligodendrocyte precursor cells can contribute to glioblastoma formation under specific oncogenic conditions. By showing that oligodendrocyte-like and oligodendrocyte lineage precursor-like cells constitute

71% of the neoplastic fraction, our results further implicate the oligodendrocyte lineage in glioblastoma pathogenesis.

To confirm the composition of the neoplastic fraction, we applied deconvolution to GIC-enriched cells from 20 glioblastoma tumors. The neoplastic fraction constituted ~90%, higher than in bulk tumors (70%), but not 100%, likely due to CD133+ purification yielding an enriched but not entirely pure population of CD133+ cells.[43] We hypothesize that the observed relative high proportion of endothelial cells may be attributed to glioblastoma's extensive vasculature, with their presence potentially reflecting contamination from tumor-associated vasculature.

Significant differences in cell-type proportions were observed among glioblastoma subclasses GBM_RTK_I, GBM_RTK_II, and GBM_MES, including variations in TAM and immune cell fractions (e.g., B cells, CD4+ T cells, NK cells, CD8+ T cells, and neutrophils). These findings are consistent with transcriptomic and immunohistochemistry analyses, which report increased tumor-infiltrating lymphocytes and TAM in the Mesenchymal GBM subtype compared to non-Mesenchymal subtypes.[14,44-46] Our results extend these subclassifications by revealing distinct differences in cellular composition detected through methylation-based deconvolution.

A key discovery in our study is that glioblastoma's cellular composition significantly influences patient survival. Specifically, higher astrocyte proportions were associated with improved overall survival probability,[17] while increased oligodendrocyte and OP proportions were linked to reduced OSP. This finding aligns with prior research showing that proliferative potential is highest in OPC-like and NPC-like states and lowest in AC-like states.[3,4] Furthermore, we found that the astrocyte-to-oligodendrocyte ratio emerged as a key prognostic marker. This supports the concept that the AC-like state, which has diminished tumor-initiating potential, is less proliferative compared to the NPC-like and OPC-like states, which are more tumorigenic in mouse models.[4]

Although GBM_RTK_I, GBM_RTK_II, and GBM_MES tumors exhibited distinct cellular compositions**,** these differences did not translate into significant OSP variations across subtypes, consistent with findings from other studies.[44,47]. Instead, as suggested above, survival differences were driven by the astrocyte-to-oligodendrocyte ratio**,** suggesting that intertumoral cellular heterogeneity may be a more powerful prognostic indicator than subclassification alone.

In conclusion, methylation-based deconvolution provides a robust approach to understanding glioblastoma's cellular complexity. The astrocyte-to-oligodendrocyte ratio serves as a prognostic marker of patient survival. Assessing cell type proportions from the initial tumor resections offers valuable prognostic insights, particularly in the context of standard therapies. This method complements single-cell RNA sequencing, facilitating analysis of archived or bulk DNA tumor samples and reducing tissue dissociation challenges. These advancements hold significant potential to refine patient stratification, guide personalized treatment strategies, and improve survival outcomes.

Limitations of the Study

This study has several limitations. First, the DNA methylation profiles of in vitro-differentiated neural progenitors and oligodendrocyte-lineage precursors may not fully capture the methylation landscapes present in glioblastoma tumors, potentially affecting the accuracy of our findings. Second, the survival analysis was conducted on a cohort predominantly receiving standard-of-care therapy, which may limit its applicability to patients undergoing alternative treatments or to underrepresented glioblastoma subtypes. Further research is required to evaluate the broader applicability of these results.

Required statements

Ethics: All human subject data used in this study were obtained from online databases, including The Cancer Genome Atlas (TCGA).

Funding: This research was supported by a generous gift from Leslie and Michael Gaffin.

Conflict of Interest: The authors declare no conflicts of interest.

Author Contributions:

Aviel Iluz[1,2], Nir Lavi[1,2,3], Hanna Charbit[1,2], Mijal Gutreiman[1,2], Masha Idelson[4], Debora Steiner[4], Etti Ben-Shushan[4], Aviad Zick[5], Amir Eden[6], Anat Mordechai[1,5], Moscovici Samuel[7], Yakov Fellig[8], Alexander Lossos[1,2,5], Joshua Moss[5], Benjamin E. Reubinoff[4], Iris Lavon*[1,2]

Conceptualization, A.I., A.E., J.M, I.L.; methodology, A.I., N.L., J.M, BE. R, M.I, D.S, E. BS, I.L. MS; software, A.I., N.L, J.M.; validation, A.I., J.M. M.I, A.Z. I.L.; formal analysis, A.I., I.L.; investigation, A.I, I.L.; resources, A.I., H.C, M.G., M.I, D.S., EB.S, M.S, J.M, BE, R, I.L.; data curation, A.I., N.L., J.M. BE. R, M.I, D.S; writing—original draft preparation, A.I., I.L.; writing-review and editing, A.I., N.L, J.M., I.L., A.L, A.M; Y.F. visualization, A.I., N.L., J.M., A.L, A.M, I.L.; supervision, I.L.; project administration, I.L.; funding acquisition, I.L. All authors have read and agreed to the published version of the manuscript.

Data Availability: Reference atlases used for deconvolution analysis and clinical data for the survival analysis cohort are available in the supplementary data.

Acknowledgments

The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

We express our sincere gratitude to Andreas von Deimling and his laboratory for performing the EPIC methylation profiling of in-vitro purified neural progenitor [28] and oligodendrocyte precursor (OP) cells, which greatly contributed to this research.

We also extend our heartfelt thanks to Leslie and Michael Gaffin for their generous support.

References

1. Ostrom QT, Bauchet L, Davis FG, et al. The epidemiology of glioma in adults: a "state of the science" review. *Neuro Oncol.* 2014; 16(7):896-913.

2. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (New York, N.Y.).* 2014; 344(6190):1396-1401.

3. Neftel C, Laffy J, Filbin MG, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell.* 2019; 178(4):835-849 e821.

4. Suvà ML, Tirosh I. The Glioma Stem Cell Model in the Era of Single-Cell Genomics. *Cancer cell.* 2020; 37(5):630-636.

5. Sharma P, Aaroe A, Liang J, Puduvalli VK. Tumor microenvironment in glioblastoma: Current and emerging concepts. *Neurooncol Adv.* 2023; 5(1):vdad009.

6. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature.* 2018; 555(7697):469-474.

7. Capper D, Stichel D, Sahm F, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathol.* 2018; 136(2):181-210.

8. Johnson KC, Anderson KJ, Courtois ET, et al. Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. *Nat Genet.* 2021; 53(10):1456-1468.

9. Verburg N, Barthel FP, Anderson KJ, et al. Spatial concordance of DNA methylation classification in diffuse glioma. *Neuro Oncol.* 2021; 23(12):2054-2065.

10. Wenger A, Ferreyra Vega S, Kling T, Bontell TO, Jakola AS, Carén H. Intratumor DNA methylation heterogeneity in glioblastoma: implications for DNA methylation-based classification. *Neuro Oncol.* 2019; 21(5):616-627.

11. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. *Nature.* 2013; 501(7467):355-364.

12. Caruso FP, Garofano L, D'Angelo F, et al. A map of tumor-host interactions in glioma at single-cell resolution. *Gigascience.* 2020; 9(10).

13. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009; 10(5):459-466.

14. Singh O, Pratt D, Aldape K. Immune cell deconvolution of bulk DNA methylation data reveals an association with methylation class, key somatic alterations, and cell state in glial/glioneuronal tumors. *Acta Neuropathol Commun.* 2021; 9(1):148.

15. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet.* 2018; 392(10149):777-786.

16. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. *Nat Struct Mol Biol.* 2013; 20(3):274-281.

17. Fernandez AF, Assenov Y, Martin-Subero JI, et al. A DNA methylation fingerprint of 1628 human samples. *Genome Res.* 2012; 22(2):407-419.

18. Hovestadt V, Jones DT, Picelli S, et al. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature.* 2014; 510(7506):537-541.

19. Sahm F, Schrimpf D, Stichel D, et al. DNA methylation-based classification and grading system for meningioma: a multicentre, retrospective analysis. *Lancet Oncol.* 2017; 18(5):682-694.

20. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016; 131(6):803-820.

21. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021; 23(8):1231-1251.

22. Loyfer N, Magenheim J, Peretz A, et al. A DNA methylation atlas of normal human cell types. *Nature.* 2023; 613(7943):355-364.

23. Moss J, Magenheim J, Neiman D, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun.* 2018; 9(1):5068.

24. Alcantara Llaguno SR, Parada LF. Cell of origin of glioma: biological and clinical implications. *Br J Cancer.* 2016; 115(12):1445-1450.

25. Darmanis S, Sloan SA, Croote D, et al. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell Rep.* 2017; 21(5):1399-1410.

26. Yuan J, Levitin HM, Frattini V, et al. Single-cell transcriptome analysis of lineage diversity in high-grade glioma. *Genome Med.* 2018; 10(1):57.

27. Yu K, Hu Y, Wu F, et al. Surveying brain tumor heterogeneity by single-cell RNA-sequencing of multi-sector biopsies. *Natl Sci Rev.* 2020; 7(8):1306-1318.

28. Davenport CF, Scheithauer T, Dunst A, et al. Genome-Wide Methylation Mapping Using Nanopore Sequencing Technology Identifies Novel Tumor Suppressor Genes in Hepatocellular Carcinoma. *Int J Mol Sci.* 2021; 22(8).

29. Nishri Y, Hampton D, Ben-Shushan E, et al. Continuous Immune-Modulatory Effects of Human Olig2+ Precursor Cells Attenuating a Chronic-Active Model of Multiple Sclerosis. *Mol Neurobiol.* 2020; 57(2):1021-1034.

30. Tannenbaum SE, Turetsky TT, Singer O, et al. Derivation of xeno-free and GMP-grade human embryonic stem cells--platforms for future clinical applications. *PLoS One.* 2012; 7(6):e35325.

31. Euskirchen P, Bielle F, Labreche K, et al. Same-day genomic and epigenomic diagnosis of brain tumors using real-time nanopore sequencing. *Acta Neuropathol.* 2017; 134(5):691-703.

32. Kuschel LP, Hench J, Frank S, et al. Robust methylation-based classification of brain tumors using nanopore sequencing. *medRxiv.* 2021:2021.2003.2006.21252627.

33. Fischer B SM, Pau G. rhdf5: R Interface to HDF5, R Package version 2.42.0. <https://github.com/grimbough/rhdf5>. 2.42.0 ed2022.

34. Vinel C, Rosser G, Guglielmi L, et al. Comparative epigenetic analysis of tumour initiating cells and syngeneic EPSC-derived neural stem cells in glioblastoma. *Nat Commun.* 2021; 12(1):6130.

35. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol.* 2012; 30(5):413-421.

36. Alboukadel K, Marcin K, Przemyslaw B. *survminer: Drawing Survival Curves using 'ggplot2'*. R package version 0.5.02024.

37. Shahar T, Granit A, Zrihan D, et al. Expression level of miRNAs on chromosome 14q32.31 region correlates with tumor aggressiveness and survival of glioblastoma patients. *J Neurooncol.* 2016; 130(3):413-422.

38. Johann PD, Jäger N, Pfister SM, Sill M. RF_Purify: a novel tool for comprehensive analysis of tumor-purity in methylation array data based on random forest regression. *BMC Bioinformatics.* 2019; 20(1):428.

39. Greenwald AC, Darnell NG, Hoefflin R, et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell.* 2024; 187(10):2485-2501.e2426.

40. Liu C, Sage JC, Miller MR, et al. Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell.* 2011; 146(2):209-221.

41. Persson AI, Petritsch C, Swartling FJ, et al. Non-stem cell origin for oligodendroglioma. *Cancer cell.* 2010; 18(6):669-682.

42. Ligon KL, Alberta JA, Kho AT, et al. The oligodendroglial lineage marker OLIG2 is universally expressed in diffuse gliomas. *J Neuropathol Exp Neurol.* 2004; 63(5):499-509.

43. Singh SK, Clarke ID, Terasaki M, et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res.* 2003; 63(18):5821-5828.

44. Wang Q, Hu B, Hu X, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer cell.* 2017; 32(1):42-56.e46.

45. Snuderl M. Molecular classification and deconvolution of the immune microenvironment in glioblastoma. *Neuro-Oncology.* 2021; 23(2):175-176.

46. Kaffes I, Szulzewsky F, Chen Z, et al. Human Mesenchymal glioblastomas are characterized by an increased immune cell presence compared to Proneural and Classical tumors. *Oncoimmunology.* 2019; 8(11):e1655360.

47. Drexler R, Schüller U, Eckhardt A, et al. DNA methylation subclasses predict the benefit from gross total tumor resection in IDH-wildtype glioblastoma patients. *Neuro Oncol.* 2023; 25(2):315-325.

Supplementary figures

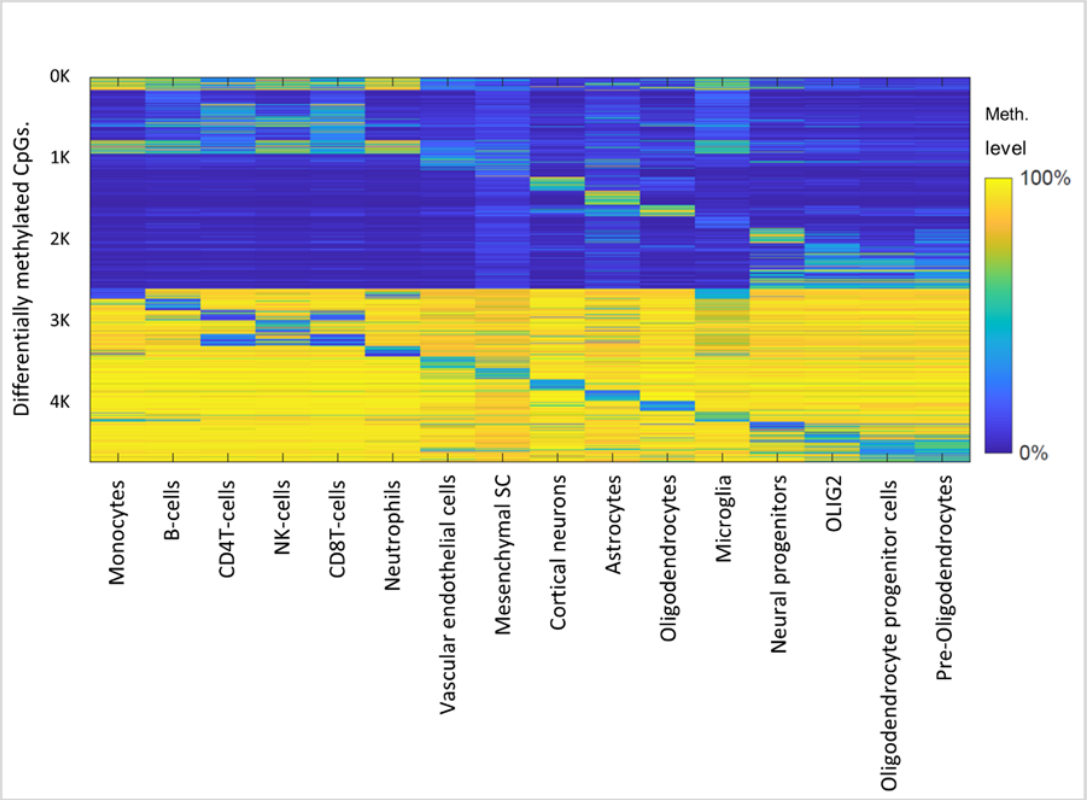

Figure S1. Reference atlas for glioblastoma deconvolution and cell-type identification using methylation signatures.

The methylation atlas comprises 16 tissues/cell types (columns) across 4712 CpGs sites (rows) that are located in 2817 genomic blocks of 500 bp each. Feature selection was carried out as described in a previous study.22 Specifically, the top 100 uniquely hypermethylated CpGs and hypomethylated CpGs for each cell type, yielding a total of 3200 tissue-specific CpGs, with neighboring CpGs within 50 bp were also included.

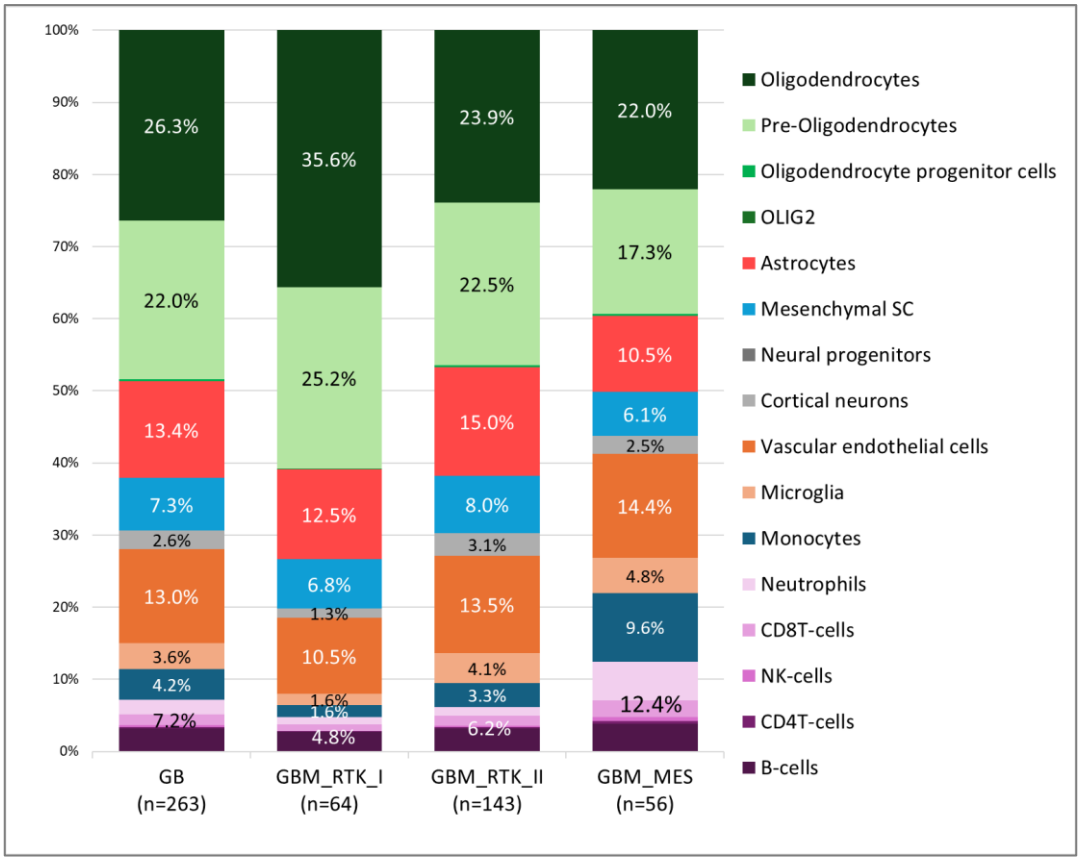

Figure S2. Methylation-based deconvolution of glioblastoma.

Proportions of normal cell types derived from the deconvolution of methylation profiles of 263 glioblastoma IDHwt samples based on the atlas of 16 cell types. B-cells, CD4 T-cells, NK cells, CD8 T-cells, and neutrophils (immune cells) are shaded in purple, with their total percentages indicating their aggregated proportion. NP, Neural progenitors; OLIG2, OLIG2-positive progenitor cells; OPC, oligodendrocyte progenitor cells; pre-oligodendrocytes, O4+ pre-oligodendrocytes.

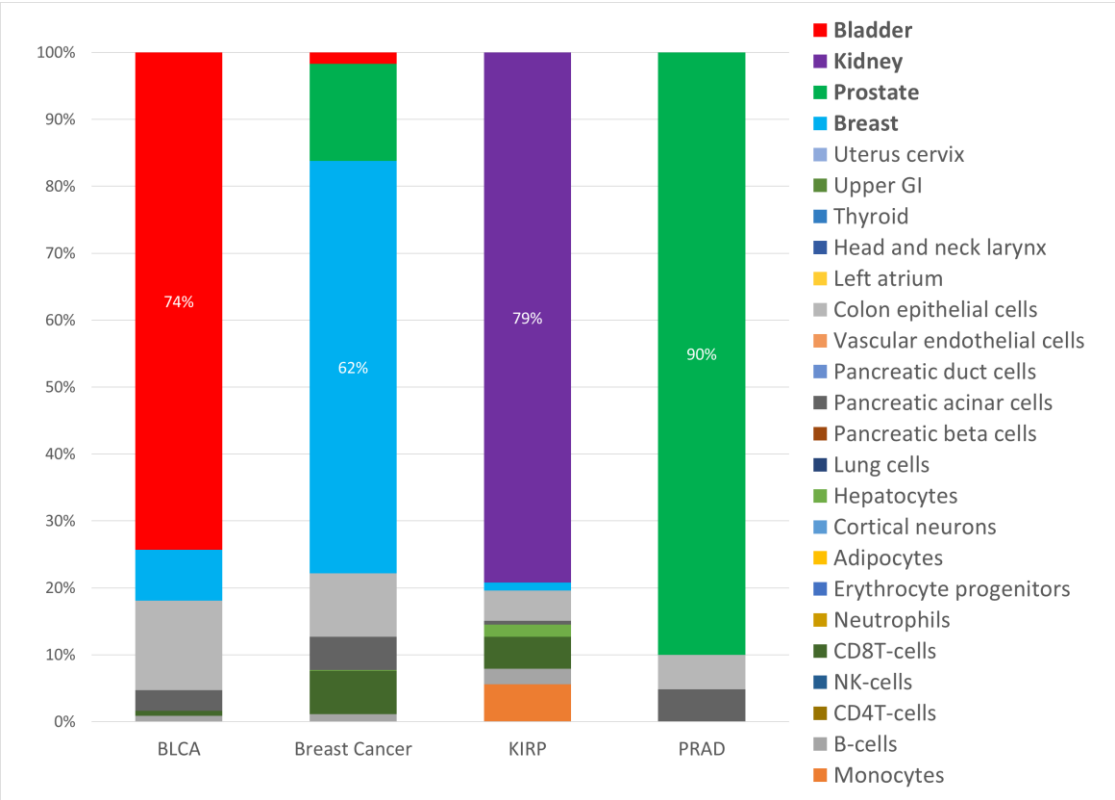

Figure S3. Methylation-based deconvolution of non-glial tumors.

Proportions of normal cell types derived from the deconvolution of methylation profiles from four different non-glial tumors (bladder urothelial carcinoma (BLCA), breast cancer, kidney renal papillary cell carcinoma (KIRP) and prostate adenocarcinoma (PRAD)).

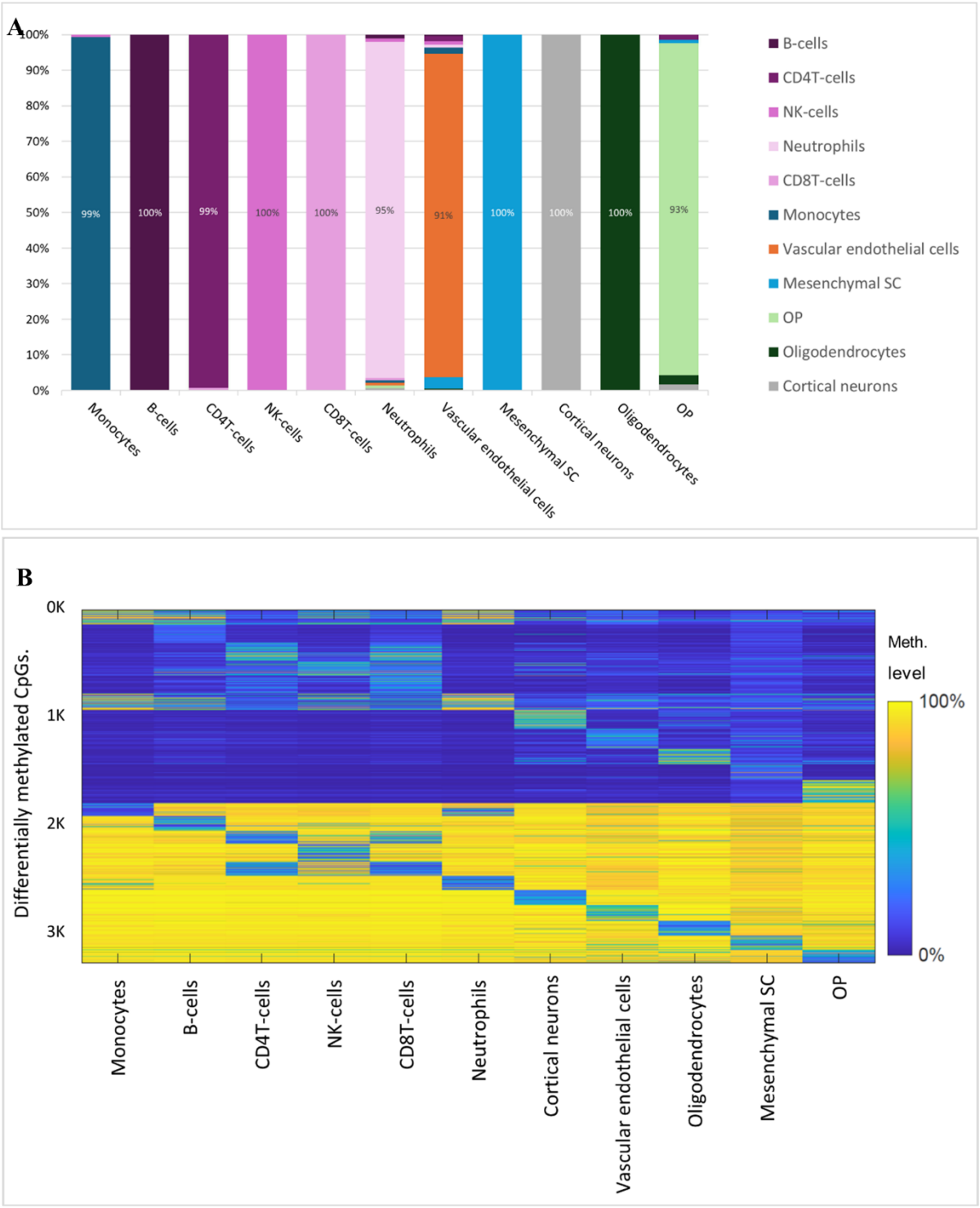

Figure S4: Methylation-based deconvolution of normal cell type samples

(A) Proportions of normal cell types derived from the deconvolution of methylation profiles of various normal cell types (n=1 per sample), using a reference atlas derived from other normal samples. (B) The reference atlas used for deconvolution includes 11 tissue-specific cell types (columns) across 3,274 CpG sites (rows). (OP, oligodendrocyte-lineage precursors).

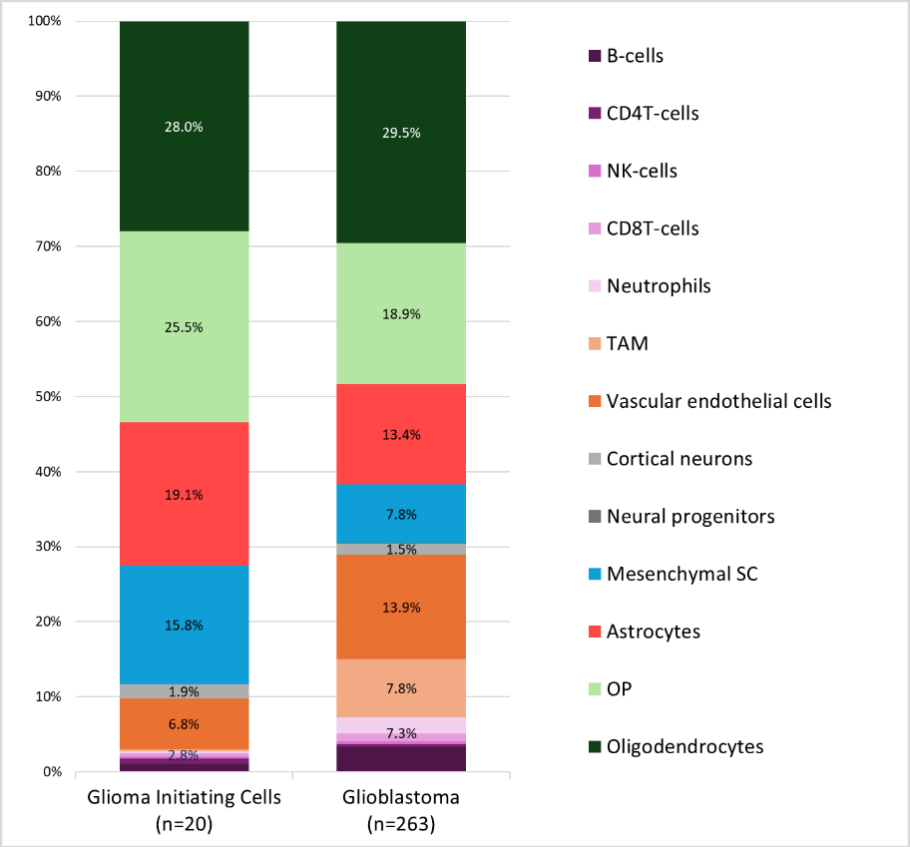

Figure S5: Methylation-based deconvolution of glioma initiating cells (GIC) samples compered to bulk tumors

Proportions of normal cell type derived from the deconvolution of methylation profiles of GIC (left bar) compared to glioblastoma (right bar): The analysis includes an average of 20 samples isolated from adult glioblastoma across three subclasses (GBM_RTK_I (n=6), GBM_RTK_II (n=6), GBM_MES (n=8)). All classifications follow Capper et al.[6] Immune cell types (B-cells, CD4 T-cells, NK cells, CD8 T-cells, and neutrophils) are shaded in purple, with their total percentages indicating their aggregated proportion. (OP, oligodendrocyte-lineage precursors).

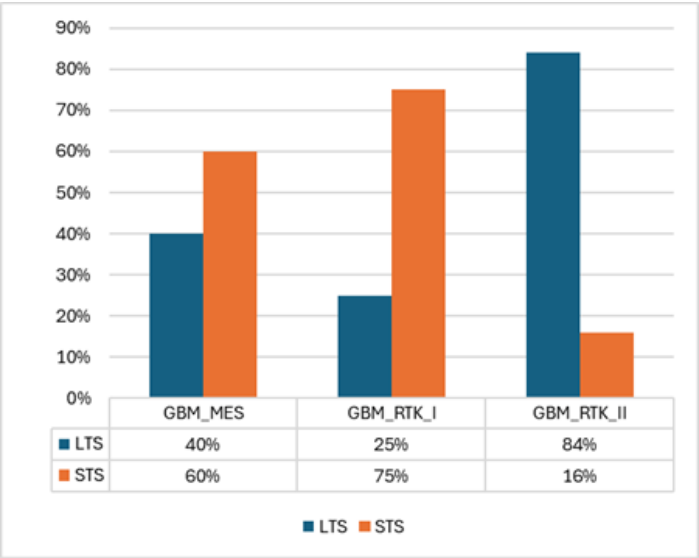

Figure S6: Distribution of glioblastoma subclasses based on the astrocyte-to-oligodendrocyte (Astro/Olig) ratio.

Histogram showing the significant proportional distribution of the three glioblastoma subclasses, GBM_RTK_I (n=24), GBM_RTK_II (n=56), and GBM_MES (n=35), across the survival outcomes determined by the astrocyte-to-oligodendrocyte proportion ratio (chi-square test, p=1.96E-07). (LTS, long-term survival; STS, short-term survival). The table below provides the patient proportions for each group.

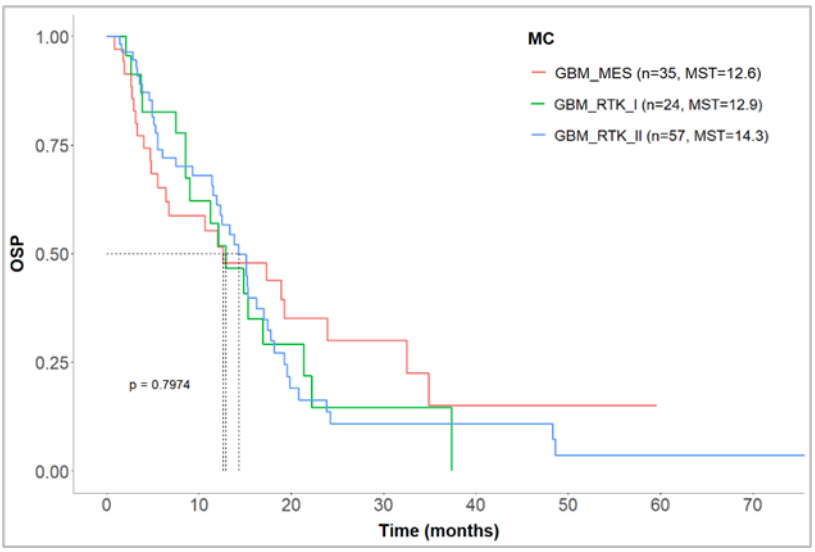

Figure S7: Survival analysis of three glioblastoma subclasses

Kaplan-Meier survival analysis of 115 TCGA glioblastoma samples classified as GBM_RTK_I (n=24), GBM_RTK_II (n=56), and GBM_MES (n=35). (MC, methylation class).